8.2 A 2Gb/s/pin 512Mb Graphics DRAM with Noise-Reduction Techniques

M. Brox[1], H. Fibranz[1], M. Kuzmenka[1], F. Lu[1], S. Mann[2], M. Markert[1], U. Möller[1], M. Plan[1], K. Schiller[1], P. Schmölz[1], P. Schrögmeier[1], A. Täuber[1], B. Weber[1], P. Mayer[1], W. Spirkl[1], H. Steffens[1], J. Weller[1]

[1]Infineon Technologies, Munich, Germany
[2]Infineon Technologies, Cary, NC

To address the bandwidth demand of today's multimedia applications, 32b-wide 256Mb graphics DRAMs have been pushed to 1.6Gb/s [1, 2]. In this paper, a 512Mb DRAM operating up to 2Gb/s is presented. To achieve this data rate, noise-reduction techniques are applied to improve IO-behavior and internal signaling. IO-behavior is influenced by simultaneous switching (SS) both for outputs (SSO) as well as for inputs (SSI). Internal signaling performance depends on cross-coupling noise of the on-chip busses.

Due to inductive noise, the internal voltages (V_{DDQi} ...) become time-dependent and differ from the static external supply voltages (V_{DDQ} ...). For SSO, a noise-sensitive transition is introduced at the interface between V_{DD} and V_{DDQ} domains within the pad driver. During any pattern, V_{SSQi} is ringing against V_{SSi} resulting in a time-dependent voltage V_{NOISE}=V_{SSQi}-V_{SSi}. The effect of noise on the current of the pad driver can be eliminated by driving the gate from the V_{DDQ} domain [3]. However, the propagation delay through the pre-driver still varies depending on V_{NOISE}. Figure 8.2.1 shows the simulated jitter at the transition for a random data sequence. To limit the effect of SSO, one has to start from the observation that the on-chip decoupling capacitance in the V_{DD} and V_{DDQ} domains guarantees

$$V_{DDi} - V_{SSi} = v_{dd}$$
$$V_{DDQi} - V_{SSQi} = v_{dd}$$

($v_{dd} = V_{DD} - V_{SS}$ is the static supply-voltage difference). Then, voltage differences across the domain can be written as,

$$V_{DDi} - V_{SSQi} = v_{dd} - V_{NOISE}$$
$$V_{DDQi} - V_{SSi} = v_{dd} + V_{NOISE}$$

e.g., if V_{SSQi} bounces positive relative to V_{SSi}, NMOS propagation at the domain-transition slows down (decrease of V_{DDi}-V_{SSQi}), whereas PMOS propagation speeds up by nearly the same amount. As a consequence, the sensitivity of the delay on V_{NOISE} can be strongly reduced by averaging an NMOS and a PMOS path (Fig. 8.2.1). The transition occurs for two branches in parallel such that a rising edge and a falling edge are crossing the domain. Independent of V_{NOISE}, one edge is accelerated and one slowed down and, at the averaging output, delay and acceleration compensate. In the example, the total propagation delay of the averaging design shows only one third of the variation of the uncontrolled transition in a conventional design.

As all high-speed DRAMs are using on-die-termination (ODT), SSI is equally important. For example, the GDDR3 standard requires ODT of 60Ω driven by an external 40Ω source. Figure 8.2.2 shows the LC-equivalent circuit of the graphics-DRAM decoupling including package inductance and a parasitic capacitance CX coupling V_{DDQ} to V_{SS}. Main contributors to CX are the n-wells of the pFETs of the pad drivers and parasitic capacitances in the deep-trench-capacitors used for decoupling of the V_{DDQ} domain. The parasitic capacitance CX is small compared to the decoupling capacitances C_{DD} and C_{DDQ} and introduces a resonance at:

$$\omega = \frac{1}{\sqrt{CX \times (L_{DD}/2 + L_{DDQ}/2)}}$$

For typical parasitic L and C values, the resonance falls into the 500MHz range and can easily be excited by a corresponding frequency component in the incoming data. Resonant oscillation on V_{SSi} is demonstrated in the simulation and measurement as shown in Fig. 8.2.2 (simulation includes realistic series resistors). Oscillation of V_{SSi} causes the received value of the write-strobe signal (voltage difference between write data strobe (WDQS) and V_{SSi} at the pads) to exhibit slope-reversals in the measurement which lead to a fail if illegal strobe-transitions are detected. As the oscillation on V_{SSi} is out of phase by 180° to V_{SSQi}, a measure to eliminate the resonance is to insert resistive damping elements between V_{SSi} and V_{SSQi}. These are implemented as NMOS-devices to be able to program the resistance dependent on the logic state of the chip as either always on or on during write. Experimentally, switches can be turned completely off to demonstrate the sensitivity of the input level on SSI. Figure 8.2.3 shows an input level versus frequency measurement where a comparison of activated and de-activated noise-suppressing elements is shown. At the resonance of 530MHz, the WDQS-level degrades to 420mV without noise-suppression whereas an input-level of better than 100mV can be achieved with activated noise-suppression.

Figure 8.2.4 shows the floorplan of the chip. Large circuits for generation of the core voltages are moved along the array to allow flexible placement of timing-critical circuitry. For the data (DQ) receivers, a variant of the matching scheme of [1] is employed, where each DQ byte is broken into two nibbles of 4b. However, to minimize delay for the on-chip write-strobe, two separate WDQS receivers are connected to the pad and placed in the center of its 4 associated DQ. Length matching for WDQS and DQ between pads and receiver guarantees best data-capture timing. Also, all driven on-chip wires are short such that complex load-balancing as in [1] can be avoided. In this approach, the WDQS pad is loaded with twice the wiring capacitance to the pad which can be compensated by adapting the load replica for the pad driver on the input-only WDQS.

Access-time is balanced independent of the bank through redistribution circuits at the center-point of each quadrant and double byte. However, coupling noise within the data bus is expected to spread read access time. As the data bus is a significant contributor to die-area for this ×32 4b pre-fetch device, fully shielding the bus is not an option. In addition, the memory core needs to operate at 500MHz. Therefore, time multiplexing a unidirectional bus to be able to shield read-by-write lines [4] is not possible. To minimize coupling effects between the data lines, an inverting driving scheme is proposed. As shown in Fig. 8.2.5, any second line is re-driven in the inverted state such that worst-case line-to-line coupling is not occurring over the full length of the bus. This method is similar to the shifted driver placement [5], however, it can be applied when the placement is defined by the architecture. The simulation compares pattern-dependent spread of propagation delay in a classical non-inverting driving scheme to the inverted driving scheme demonstrating an improvement of over 500ps for a core cycle time of 2ns.

The 16M×32b graphics DRAM is manufactured in a conventional 0.11µm DRAM process. Figure 8.2.6 demonstrates 2Gb/s operation at a read-latency of 11 cycles. Figure 8.2.7 shows a die micrograph.

References:

[1] H.Y. Song, et al. "A 1.2Gb/s/pin Double Data Rate SDRAM with On-Die-Termination," *ISSCC Dig. Tech. Papers*, pp. 314-315, Feb., 2003.
[2] S.B. Lee, et al. "A 1.6Gb/s/pin Double-Data-Rate SDRAM with Wave-Pipelined CAS Latency Control," *ISSCC Dig. Tech. Papers*, pp. 210-211, Feb., 2004.
[3] J. B. Lee, et al. "Digitally-controlled DLL and I/O circuits for 500Mb/s/pin ×16 DDR SDRAM," *ISSCC Dig. Tech. Papers*, pp. 68-69, Feb., 2001.
[4] C. Yoo, et al., "A 1.8-V 700-Mb/s/pin 512-Mb DDR-II SDRAM With On-Die Termination and Off-Chip Driver Calibration," *IEEE J. Solid-State Circuits*, vol. 39, no. 6, pp. 941-951, Jun., 2004.
[5] S. Naffziger, Tutorial, ISSCC 1999.

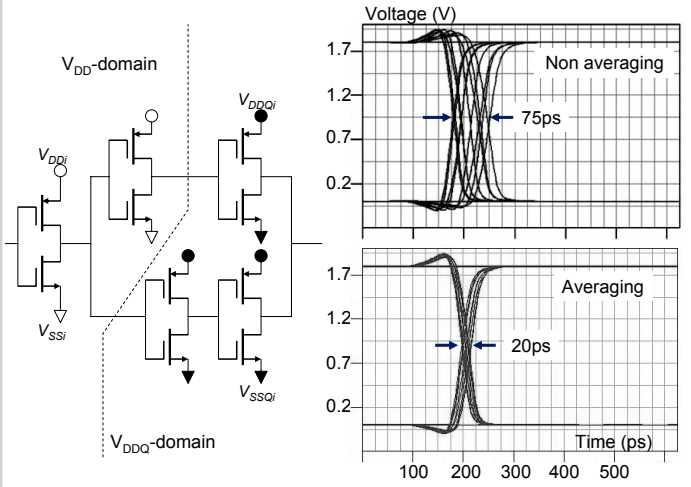

Figure 8.2.1: Averaging at domain transition and simulation of noise without and with averaging. PRBS with sinusoidal noise of Vpp=350mV at f=1GHz on V_{SSQi}.

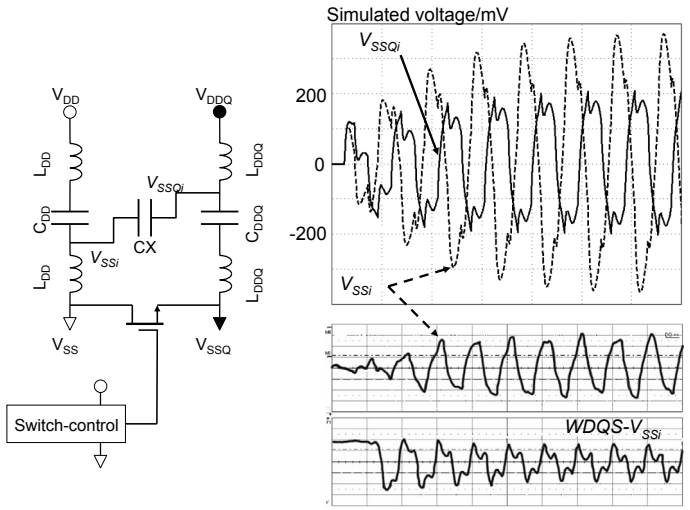

Figure 8.2.2: LC-model of power-supply network with simulation at the resonance frequency 530MHz. Measurement of V_{SSi} and on-die difference (WDQS-V_{SSi}).

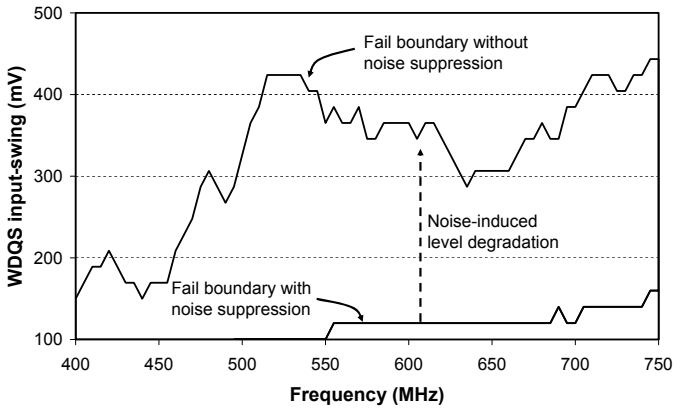

Figure 8.2.3: Required WDQS input level swing with and without noise suppression. DQ input level swing: 570mV. Clock-like input pattern on all DQ.

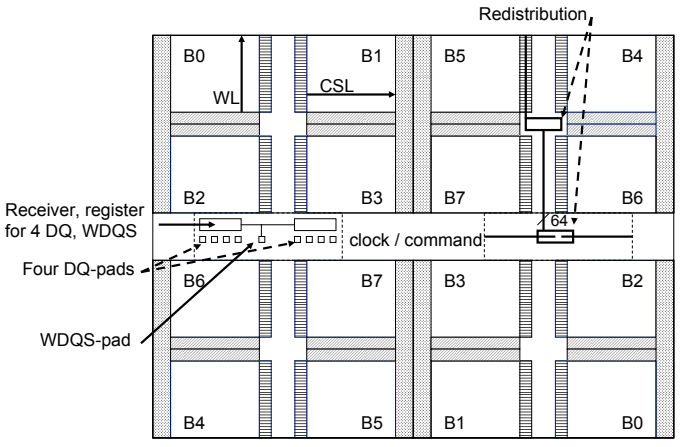

Figure 8.2.4: Floor-plan. Bn: half bank; WL: wordline; CSL: column select line.

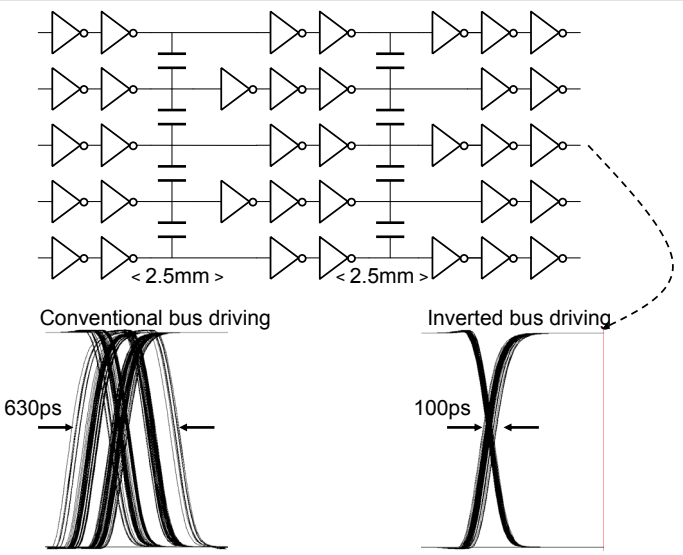

Figure 8.2.5: Waveform at the end of bus for conventional and inverted drive.

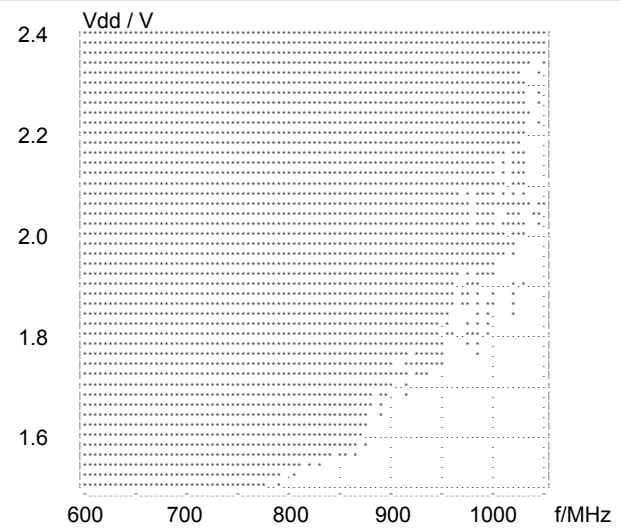

Figure 8.2.6: Frequency versus supply voltage plot at T=25C. Fail soccurring at V_{dd}-independent frequencies are caused by jitter tolerances within the test system.

Continued on Page 645

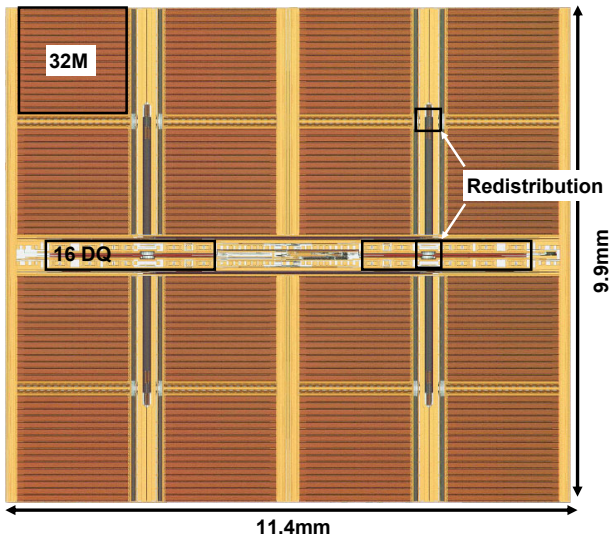

Figure 8.2.7: Die micrograph.

